NOTICE OF EXPRESS MAILING

Express Mail Mailing Label Number: **EL740531258US**

Date of Deposit with USPS: **April 10, 2001**

Person making Deposit: **Daniel Thatcher**

APPLICATION FOR LETTERS PATENT

for

INTEGRATED CIRCUIT HAVING SELF-ALIGNED CVD-TUNGSTEN/TITANIUM CONTACT PLUGS STRAPPED WITH METAL INTERCONNECT AND METHOD OF MANUFACTURE

Inventors:
Kirk D. Prall
Howard E. Rhodes
Sujit Sharan
Gurtej S. Sandhu
Philip J. Ireland

Attorney:
James R. Duzan
Registration No. 28,393
TRASKBRITT
P.O. Box 2550
Salt Lake City, Utah 84110
(801) 532-1922

INTEGRATED CIRCUIT HAVING SELF-ALIGNED
CVD-TUNGSTEN/TITANIUM CONTACT PLUGS STRAPPED WITH
METAL INTERCONNECT AND METHOD OF MANUFACTURE

BACKGROUND OF THE INVENTION

[0001] <u>Cross Reference to Related Application</u>: This application is a divisional of application Serial No. 09/447,981, filed November 23, 1999, pending, which is a continuation of application Serial No. 08/994,950, filed December 19, 1997, now U.S. Patent No. 5,990,021, issued November 23, 1999, which is a divisional of application Serial No. 08/604,344, filed February 21, 1996, abandoned.

[0002] <u>Field of the Invention</u>: This invention relates to integrated circuits and, more particularly, to processes for fabricating high aspect ratio contact plugs, which are particularly useful in dynamic random access memories having stacked capacitors.

[0003] <u>State of the Art</u>: The business of producing dynamic random access memory (DRAM) devices is a very competitive, high-volume business. Process efficiency and manufacturability, as well as product quality, reliability, and performance are the key factors that determine the economic success of such a venture.

[0004] Each cell within a DRAM device, an individually-addressable location for storing a single bit of digital data, is comprised of two main components: a field-effect access transistor and a capacitor. Each new generation of DRAM devices generally has an integration level that is four times that of the generation which it replaced. Such a quadrupling of the number of devices per chip is always accompanied by a decrease in device geometries, and often by a decrease in operating voltages. As device geometries and operating voltages are decreased, the DRAM designer is faced with the difficult task of providing reliable, low-resistance electrical interconnections on the chip, increasing device operating speeds, and maintaining cell capacitance at an acceptable level. This must be accomplished without resorting to processes that reduce product yield or that markedly increase the number of masking and deposition steps in the production process.

[0005] During integrated circuit manufacture, it is common for contact openings to be etched through an insulative layer down to a diffusion region to which electrical contact is to be

made. In a conventional contact formation process, titanium metal is sputtered over the wafer so that the exposed surface of the diffusion region is coated. The titanium metal is eventually converted to titanium silicide, thus providing an excellent conductive interface at the surface of the diffusion region. A titanium nitride barrier layer is then deposited, coating the walls and floor of the contact opening. Chemical vapor deposition of tungsten or polycrystalline silicon ("polysilicon") follows. In the case of tungsten, the titanium nitride layer provides greatly improved adhesion between the walls of the opening and the tungsten metal, in addition to preventing attack of the substrate and the formation of "wormholes" by tungsten hexafluoride, the precursor compound typically used for tungsten deposition. In the case of the polysilicon, the titanium nitride layer acts as a barrier against dopant diffusion from the polysilicon layer into the diffusion region.

[0006] Several factors have been responsible for an increase in contact aspect ratios. Given the need to maintain cell capacitance as the cell area dedicated to each cell on a wafer is shrunk, it has been necessary to employ three-dimensional capacitors for DRAMs of the 4-megabit generation and beyond. Although trench capacitors fabricated in the substrate have been utilized by several manufacturers, most have chosen to employ stacked capacitors (i.e., stacked above the substrate). Because stacked capacitors normally employ vertically-oriented capacitive layers, the stack height (i.e., the vertical distance from the substrate to the top of an insulating layer overlying the capacitive layers) has been increasing significantly with each new generation of stacked capacitor DRAMs. The increasing stack height coupled with the decreased area on the substrate to which electrical contact must be made has mandated an increase in contact aspect ratios (contact aspect ratio being the ratio of contact depth to contact opening width).

[0007] As decreasing device geometries shrink, the substrate area available for making contact to transistor source/drain regions (i.e., the contact landing pad), circuit designers are utilizing self-aligned contacts with increasing frequency. A self-aligned contact is one in which the contact plug is immediately adjacent a transistor gate spacer. With a self-aligned contact structure, the pitch between transistor gates can be narrowed, as the entire landing pad is utilized for making contact to the substrate. Thus, for a given pitch, contact resistance is minimized and reliability is enhanced. A self-aligned contact requires that the gate electrode be coated with a

3

dielectric material that is largely unaffected by the anisotropic plasma contact etch through the interlevel dielectric layer.

[0008]   The use of self-aligned contacts encourages the use of minimum-photolithographic-pitch layouts. A minimum-pitch layout coupled with the high-aspect ratio contacts that result in DRAM devices having stacked capacitors is not a good candidate for conventional contact formation, as contact between the small substrate landing pad and the contact conductive material must be particularly sound. For the 256-megabit generation of DRAMs, contacts having contact aspect ratios of greater than 5:1 will likely be the norm. With contact openings having such high aspect ratios, sputtering is ineffective at coating the bottom of the contact opening with titanium metal, as it tends to deposit primarily near the mouth of the opening such that the size of the opening becomes narrowed. The narrowing effect reduces the amount of titanium that can be deposited at the bottom of the opening. Since the bottom of the opening is also restricted in size by the use of a self-aligned contact, electrical contact to the substrate is typically poor. Although it is possible to collimate sputtered metal, thereby decreasing the buildup of sputtered material near the mouth of the contact opening and thereby increasing the amount of material that can be deposited at the bottom of the contact opening, deposition rates drop dramatically and deposition times correspondingly increase as collimator aspect ratios increase. At collimator height-to-width aspect ratios of only 2.5:1, deposition rates on the wafer are nearly zero, with most metal atoms being deposited on the collimator. Thus, the collimators must be cleaned with a frequency that is annoying, if not unacceptable, in a production environment. To date, the solutions proposed for this problem have required manufacturing process flows of increased complexity, which invariably tend to decrease product yield and increase product costs. For example, silicided polysilicon buried digit lines (a buried digit line is located at a level below that of the stacked capacitors) have been almost universally adopted by DRAM manufacturers as a solution to high-aspect ratio digit line contacts. The downside of using a silicided polysilicon buried digit line is that the increased resistance, as compared with that associated with a metal digit line, results in longer information access times. In addition, a buried digit line process typically requires two additional masking steps. Another solution to the problem is to use a polysilicon plug which makes contact to the substrate landing pad and a

4

superjacent metal plug which makes contact to the polysilicon plug. The metal plugs are then strapped with a low-resistance metal interconnect. The downside of such a process flow is added complexity and increased product coat.

[0009]   What is needed is a simplified DRAM process flow which takes advantage of a self-aligned contact for reduced pitch and greater chip density, and which is compatible with high-aspect-ratio contacts. Such a process and resulting structure will enhance device speed by permitting the use of a metal interconnect structure on top of the cell capacitors.

SUMMARY OF THE INVENTION

[0010]   This invention is a process for manufacturing a dynamic random access memory array. Each memory cell within the array which results from the process incorporates a stacked capacitor, a silicon nitride coated access transistor gate electrode, and a self-aligned high-aspect-ratio digit line contact having a tungsten plug which extends from the substrate to a metal interconnect structure located at a level above the stacked capacitor. The contact opening is lined with titanium metal, which is in contact with the substrate, and with titanium nitride, which is in contact with the plug. Both the titanium metal and the titanium nitride are deposited via chemical vapor deposition reactions.

[0011]   The DRAM process begins after field oxide and gate oxide formation with the patterning of word lines (i.e., the access transistor gates and gate interconnect lines) from a sandwich consisting of a lower layer of doped polysilicon, a middle layer of a refractory metal silicide (e.g., titanium silicide, tungsten silicide, cobalt silicide), and an upper layer of silicon nitride. Following word line formation, spacers are formed on the vertical edges of the word lines by anisotropically etching a subsequently deposited conformal silicon nitride layer. Following the formation of transistor source/drain regions (preferably through several implantation steps), a stacked capacitor is constructed above each source/drain which functions as a storage-node junction. During, following, or during and following the construction of the stacked capacitors, an interlevel silicon dioxide based dielectric layer such as boro-phospho-silicate glass ("BPSG") is deposited and preferably employed to planarize the array above the level of the stacked capacitors. Contact openings which are self aligned to the silicon nitride transistor spacers are

5

then etched through the BPSG layer to each of the transistor source/drain regions in the substrate which will function as an access-node junction. Each of the contact openings is lined, first with titanium metal, and later with titanium nitride. Both the titanium metal and the titanium nitride layers are deposited via chemical vapor deposition, which provides greatly improved step coverage over sputter deposition methods. During a subsequent anneal step, the titanium metal which is in contact with each of the junctions at the bottom of the contact openings is converted to titanium silicide. The contact openings are then filled with tungsten metal via a further chemical vapor deposition reaction. The tungsten which deposited on top of the interlevel dielectric layer is removed via either a chemical mechanical polishing step or an etch-back step. Metal digit line interconnects which contact the appropriate tungsten plugs are then patterned from a sputtered aluminum alloy layer.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWINGS

[0012]   FIG. 1 is a cross-sectional view of a portion of an in-process DRAM array following field oxide and gate oxide formation and the patterning of word lines;

[0013]   FIG. 2 is a cross-sectional view of the in-process DRAM array portion of FIG. 1 following the formation of access transistor source/drain regions and the formation of spacers on the word line sidewalls;

[0014]   FIG. 3 is a cross-sectional view of the in-process DRAM array portion of FIG. 2 following the formation of typical stacked capacitors and the deposition and planarization of an interlevel dielectric layer;

[0015]   FIG. 4 is a cross-sectional view of the in-process DRAM array portion of FIG. 3 following the masking of the interlevel dielectric layer with a digit line contact mask and etching of the contact openings through the interlevel dielectric layer to the underlying junctions in the substrate;

[0016]   FIG. 5 is a cross-sectional view of the in-process DRAM array portion of FIG. 4 following the sequential deposition of a titanium metal layer, a titanium nitride layer, and a tungsten layer;

[0017] FIG. 6 is a cross-sectional view of the in-process DRAM array portion of FIG. 5 following a chemical mechanical polishing step which removes the tungsten, the titanium nitride, and the titanium metal from the upper surface of the interlevel dielectric layer; and

[0018] FIG. 7 is a cross-sectional view of the in-process DRAM array portion of FIG. 6 following the deposition of a low-resistance metal layer and patterning of that metal layer to form a digit line interconnect structure.

DETAILED DESCRIPTION OF THE INVENTION

[0019] Referring now to FIG. 1, a cross-sectional view of a portion of the memory cell array of an in-process dynamic random access memory is shown. At this stage of processing, channel-stop regions 11 have been implanted in the substrate 12, field oxide regions 13 have been formed (with present technology, a modified local oxidation of silicon, or LOCOS, process is deemed to be the preferred method), a silicon dioxide gate dielectric layer 14 has been grown on the substrate, and word lines 15 have been formed (patterned) by masking and anisotropically etching a sandwich consisting of a doped polysilicon lower layer 16, a refractory metal silicide intermediate layer 17 (e.g., titanium silicide, tungsten silicide, cobalt silicide) which is employed to reduce word line resistance, and a silicon nitride upper layer 18. Anisotropic plasma etches are typically used for the etch which forms the word lines, as they provide structures with sidewalls which are substantially vertical (i.e., vertical or within several degrees of being vertical). When the term "vertical" is used in this disclosure, exact verticality is not to be implied, as perfect geometric relationships exist only in theory. The photoresist mask 19 used to pattern the word lines has not yet been stripped from the in-process circuitry. The term "word line" is used to denote the sets of serially connected field-effect transistor gates and gate interconnects in a memory array circuit. It should be understood that whenever a word line 15 traverses an active area (active areas are the regions of the substrate which are not covered by field oxide 13), it functions as the gate of a cell access transistor. Conversely, when a word line traverses a field oxide region 13, it functions as an interconnect line which couples the gates of a single row of cells within the array to a word line driver device (not shown). Anti-punch through or lightly-

7

doped source/drain regions 20 have also been implanted into those portions of the active areas which are exposed following the word line etch.

[0020] Referring now to FIG. 2, the array portion of FIG. 1 has received further processing to create source/drain regions for the cell access transistors. The centrally positioned source/drain region 21A, which is shared by both adjacent transistor gates, will function as the access-node junction, to which digit line contact will be made. The laterally located source/drain regions 21B, each of which will function as a storage-node junction, are adjacent a single transistor gate and, thus, are employed by only one memory cell. Silicon nitride spacers 22 are formed on the word line sidewalls by anisotropically etching an unmasked, conformally-deposited silicon nitride layer (not shown). The silicon nitride upper layer 18 on the upper surface of the word line conductive portions, in combination with the nitride spacers 22 on the sidewalls of the word lines, is referred to herein as the word line isolation layer. Although the formation of source/drain regions for ULSI circuitry is a complicated process, as junctions having graded dopant densities are generally required to eliminate short channel effects, this disclosure will not deal with those complexities, as they are irrelevant to the focus of the present invention. Following transistor processing, a first interlevel dielectric layer 35A is deposited.

[0021] Referring now to FIG. 3, the array portion of FIG. 2 is subjected to further processing to form stacked capacitors 31. Each capacitor 31, which is superjacent a storage-node junction 21B, comprises a lower capacitive layer (the storage-node plate) 32 which is in electrical contact with the storage-node junction 21B, an upper capacitive layer (the cell plate) 33, and a capacitor dielectric layer 34 sandwiched between the two capacitive layers. It will be noted that a portion of the upper capacitive layer 33 has been etched away above the access-node junction 21A in order to eliminate complications of having to later put a contact plug through an unrelated conductive layer and from which it must be insulated. The capacitor dielectric layer 34, on the other hand, is optionally left unetched over the access-node junction 21A. During the fabrication of capacitors 31, the first interlevel dielectric layer 35A (preferably BPSG) is employed as a mold layer in which is formed the storage node plates 32. Following the completion of capacitor fabrication, a second interlevel dielectric layer 35B (also, preferably BPSG) is deposited over the entire array and planarized (preferably via a chemical mechanical polishing step) to a

8

level above the stacked capacitors. It should be noted that once the second interlevel dielectric layer 35B is annealed, it fuses with the first interlevel dielectric layer 35A, becoming indistinguishable therefrom. In the appended claims, the two layers are referred to as a single interlevel dielectric layer.

[0022] Referring now to FIG. 4, the in-process DRAM array portion of FIG. 3 is masked with a digit line contact mask 41 and anisotropically etched to form contact openings 42 (only one shown) through the interlevel dielectric layer 35 (comprised of layers 35A and 35B) to the underlying access-node junctions 21A in the substrate. It will be noted that the lower portion of the contact opening is self-aligned to the spacers 22 on either side of the access-node junction 21A. This result will naturally follow if the word lines are formed at or near minimum photolithographic pitch. At minimum photolithographic pitch, space and feature widths are at the resolution limit of the stepper (the machine employed to expose the photoresist). Since the spacers on the sidewalls of the word lines have narrowed the minimum-pitch space between the word lines, and because the minimum space for the digit line contact mask 41 is the same as for the mask used to etch the word lines (assuming that the same type of photolithographic equipment is employed for mask exposure), the contact opening, as defined by the digit line contact mask 41 will be wider than the landing pad 43 of the access-node junction 21A, which has been narrowed by the spacers 22. In order to prevent shorting of the digit line (not yet formed) to the word lines, an anisotropic plasma etch that etches the interlevel dielectric layer 35 selectively with respect to the spacers 22 and the silicon nitride layer 18 atop the word lines is required. Anisotropic plasma etches that etch BPSG selectively (i.e., the etch rate of BPSG is so much greater than that of silicon nitride that the etching of silicon nitride is insignificant) with respect to silicon nitride are well known in the art. It should be noted that during the etch of the contact opening, it may be necessary to etch through the capacitor dielectric layer 34 if it was optionally left unetched over the access node junction 21A. If the capacitor dielectric layer comprises silicon nitride, the selectivity is decreased during an early portion of the etch so that the silicon nitride layer may be easily penetrated.

[0023] Referring now to FIG. 5, following an in situ clean step (preferably using a plasma etch) to remove the native oxide from the surface of the substrate at the bottom of the

9

contact opening, a titanium metal layer 51 is deposited on all exposed surfaces of the in-process DRAM array portion of FIG. 4 via chemical vapor deposition (CVD). The preferred precursor gases for the CVD reaction are $TiCl_4$, with a flow rate of about 30 sccm, and H_2, with a flow rate of about 100 sccm. Argon, an inert gas, is also introduced into the reaction chamber. The chamber is equipped with a shower head that is RF-powered (e.g., 400kHZ to 13.56MHz) to create a plasma. The temperature range for the wafer is 450-550° C. The chamber pressure is 5-10 torr. Thus the titanium metal layer 51 lines the interior of the contact opening 42 and covers the landing pad (the portion of the junction exposed by the contact etch) of access-node junction 21A. As the contact opening has a high aspect ratio (in this case, it is greater than 5:1), the deposition of titanium metal via a chemical vapor (CVD) reaction provides the step coverage required to adequately cover the junction at the bottom of the contact opening 42. At this stage of the process, an elevated-temperature anneal step may be performed to convert at least a portion of the titanium metal layer 51 that overlies the access-node junction 21A to a titanium silicide layer 52. Following the anneal step, a titanium nitride layer 53 is deposited on all exposed surfaces of the array via chemical vapor deposition. A preferred CVD reaction for titanium nitride is discussed in U.S. patent 5,192,589, and employs tetrakis(dimetylamido)titanium and ammonia as precursors. The excellent step coverage afforded by this chemical vapor deposition reaction results in relatively conformal deposition of the titanium nitride layer 53 within the entire contact opening 42. A variation in the process would be to perform the aforementioned anneal step following the deposition of the titanium nitride layer 53. Finally, a tungsten layer 54 is also deposited via chemical vapor deposition. CVD deposition of tungsten is already well known in the art and no further discussion is deemed necessary. The tungsten layer 54 is deposited to a thickness sufficient to completely fill the contact opening 42.

[0024] Referring now to FIG. 6, the array portion of FIG. 5 is subjected to a chemical mechanical polishing step which removes portions of the tungsten layer 54, portions of the titanium nitride layer 53, and portions of the titanium metal layer 51 which lie on the upper surface of the interlevel dielectric layer 35, leaving those portions of the tungsten layer 54, the titanium nitride layer 53 and the titanium metal layer 51 which are within the confines of the contact opening 42. A tungsten plug 61 remains in each digit line contact opening.

10

[0025]   Referring now to FIG. 7, an aluminum alloy layer is deposited on the array portion of FIG. 6, masked with a digit line interconnect mask (not shown) and etched to form a plurality of digit lines 71 (only one digit line is depicted in the figure) on the upper surface of the interlevel dielectric layer 35. Each digit line makes electrical contact to the tungsten plugs 61 which pertain to a single column of memory cells within the array.

[0026]   Although only a single embodiment of the invention (and minor variations thereof) is disclosed herein, it will be obvious to those having ordinary skill in the art of integrated circuit manufacture that changes and modifications may be made thereto without departing from the scope and the spirit of the invention as hereinafter claimed. For example, the process, though disclosed in connection with a dynamic random access memory having stacked capacitors, is applicable to any integrated circuit requiring high-aspect ratio contacts.